Genome Glossary

Interphase The period in the cell cycle when DNAis replicated in the nucleus; followed by mitosis.

Intron The DNA base sequence interrupting the protein coding sequence of a gene; this sequence is transcribed into RNA but is cut out of the message before it is translated into protein. Compare exon.

In vitro Outside a living organism.

K

Karyotype A photomicrograph of an individual's chromosomes arranged in a standard format showing the number, size, and shape of each chromosome type; used in low-resolution physical mapping to correlate gross chromosomal abnormalities with the characteristics of specific diseases.

kb See kilobase.

Kilobase (kb) Unit of length for DNA fragments equal to 1000 nucleotides.

L

Library An unordered collection of clones from a particular organism, whose relationships to each other can be established by physical mapping. Compare genomic library, arrayed library.

Locus (pl. loci) The position on a chromosome of a gene or other chromosome marker; also, the DNA at that position. The term locus is sometimes restricted to mean regions of DNA which are expressed. See gene expression.

M

Mapping See gene mapping, linkage map, physical map.

Marker An identifiable physical location on a chromosome (e.g., restriction enzyme cutting site, gene) whose inheritance can be monitored. Markers can be expressed regions of DNA (genes) or some segment of DNA with no known coding function but whose pattern of inheritance can be determined. See restriction fragment length polymorphism.

Mb See megabase.

Megabase (Mb) Unit of length for DNA fragments equal to 1 million nucleotides and roughly equal to 1 cM.

Messenger RNA (mRNA) RNA that serves as a template for protein synthesis. See genetic code.

mRNA See messenger RNA.

Multifactorial or multigenic disorder See polygenic disorder.

Multiplexing An approach to many types of laboratory procedures where several samples are pooled and handled simultaneously, greatly increasing processing speed.

Mutation Any heritable change in DNA sequence. Compare polymorphism.

N

Nitrogenous base A nitrogen-containing molecule having the chemical properties of a base.

Nucleic acid A large molecule composed of nucleo-tide sub-units.

Nucleotide A subunit of DNA or RNA consisting of a nitrogenous base (adenine, guanine, thymine, or cytosine in DNA; adenine, guanine, uracil, or cytosine in RNA), a phosphate molecule, and a sugar molecule (deoxyribose in DNA and ribose in RNA). Thousands of nucleotides are linked to form a DNA or RNA molecule. See DNA, base pair, RNA.

O

Nucleus The cellular organelle in eukaryotes that contains the genetic material.

Overlapping clones See Contig.

P

P1-derived artificial chromosome (PAC) A vector used to clone DNA fragments (100- to 300-kb insert size; average, 150 kb) in *Escherichia coli* cells. Based on bacteriophage (a virus) P1 genome. Compare cloning vector.

PAC See P1-derived artificial chromosome.

PCR See polymerase chain reaction.

Phage A virus for which the natural host is a bacterial cell.

Physical map A map of the locations of identifiable landmarks on DNA (e.g., restriction enzyme cutting sites, genes), regardless of inheritance. For the human genome, the lowest-resolution physical map is the banding patterns on the 24 chromosomes; the highest resolution map would be the complete nucleotide sequence of the chromosomes.

Plasmid Autonomously replicating, extrachromosomal circular DNA molecules, distinct from the normal bacterial genome and nonessential for cell survival under nonselective conditions. Some plasmids are capable of integrating into the host genome. A number of artificially constructed plasmids are used as cloning vectors.

Polymerase chain reaction (PCR) A method for amplifying a DNA base sequence using a heat stable polymerase and two 20-base primers, one complementary to the (+) strand at one end of the sequence to be amplified and the other complementary to the (-) strand at the other end. Because the newly synthesized DNA strands can subsequently serve as additional templates for the same primer sequences, successive rounds of primer annealing, strand elongation, and dissociation produce rapid and highly specific amplification of the desired sequence. PCR also can be used to detect the existence of the defined sequence in a DNA sample.

Polymerase, DNA or RNA Enzymes that catalyze the synthesis of nucleic acids on preexisting nucleic acid templates, assembling RNA from ribonucleotides or DNA from deoxyribonucleotides.

Polymorphism Difference in DNA sequence among individuals. Genetic variations occurring in more than 1% of a population would be considered useful polymorphisms for genetic linkage analysis. Compare mutation.

Primer Short preexisting single stranded polynucleotide chain to which new deoxyribonucleotides can be added using DNA polymerase.

Probe Single-stranded DNA or RNA molecules of specific base sequence, labeled either radioactively or immunologically, used to detect the complementary base sequence by hybridization.

Prokaryote Cell or organism lacking a membrane-bound, structurally discrete nucleus and other sub-cellular compartments. Bacteria are prokaryotes. Compare eukaryote. See chromosome.

Promoter A site on DNA to which RNA polymerase will bind and initiate transcription.

Protein A large molecule composed of one or more chains of amino acids in a specific order; the order is deter-mined by the base sequence of nucleotides in the gene coding for the protein. Proteins are required for the structure, function, and regulation of the body's cells, tissues, and organs. Each protein has a unique function. Examples are hormones, enzymes, and antibodies.

Purine A nitrogen-containing, basic compound that occurs in nucleic acids. The purines in DNA and RNA are adenine and guanine.

Pyrimidine A nitrogen-containing, basic compound that occurs in nucleic acids. The pyrimidines in DNA are cytosine and thymine; in RNA, cytosine and uracil.

R

Recombinant clone Clone containing recombinant DNA molecules. See recombinant DNA technology.

Recombinant DNA molecules A combination of DNA molecules of different origin that are joined using recombinant DNA technologies.

Recombinant DNA technology Procedure used to join together DNA segments in a cell-free system (an environment outside a cell or organism). Under appropriate conditions, a recombinant DNA molecule can enter a cell and replicate there, either autonomously or after it has become integrated into a cellular chromosome.

Recombination The process by which progeny derive a combination of genes different from that of either parent. In higher organisms, this can occur by crossing over.

Regulatory region or sequence A sequence that controls gene expression.

Resolution Degree of molecular detail on a physical map of DNA, ranging from low to high.

Restriction enzyme, endonuclease A protein that recognizes specific, short nucleotide sequences and cuts DNA at those sites. Bacteria contain over 400 such enzymes that recognize and cut over 100 different DNA sequences. See restriction enzyme cutting site.

Restriction enzyme cutting site A specific nucleotide sequence of DNA at which a particular restriction enzyme cuts the DNA. Some sites occur frequently in DNA (e.g., every several hundred base pairs), others much less frequently (rare cutter; e.g.,every 10,000 base pairs).

Restriction fragment length polymorphism (RFLP) Variations between individuals in DNA fragment sizes cut by specific restriction enzymes; polymorphic sequences that result in RFLPs are used as markers on both physical maps and genetic linkage maps. RFLPs are usually caused by mutation at a cutting site. See marker.

RFLP See restriction fragment length polymorphism.

Ribonucleic acid (RNA) A chemical found in the nucleus and cytoplasm of cells; it plays an important role in protein synthesis and other chemical activities of the cell. The structure of RNA is similar to that of DNA. There are several classes of RNA molecules, including messenger RNA, transfer RNA, ribosomal RNA, and other small RNAs, each serving a different purpose.

Ribosomal RNA (rRNA) A class of RNA found in the ribosomes of cells.

Ribosomes Small cellular components composed of specialized ribosomal RNA and protein; site of protein synthesis. See ribonucleic acid (RNA).

RNA See ribonucleic acid.

S

Sequence See base sequence.

Sequence tagged site (STS) Short (200 to 500 base pairs) DNA *sequence* that has a single occurrence in the human genome and whose location and base sequence is known. Detectable by polymerase chain reaction, STSs are useful for localizing and orienting the mapping and sequence data reported from many different laboratories and serve as landmarks on the developing physical map of the human genome. Expressed sequence tags (ESTs) are STSs derived from cDNAs.

Sequencing Determination of the order of nucleotides (base sequences) in a DNA or RNA molecule or the order of amino acids in a protein.

Shotgun method Cloning of DNA fragments randomly generated from a genome. See library, genomic library.

Somatic cell Any cell in the body except gametes and their precursors.

STS See sequence tagged site.

T

Tandem repeat sequences Multiple copies of the same base sequence on a chromosome; used as a marker in physical mapping.

Telomere The end of a chromosome. This specialized structure is involved in the replication and stability of linear DNA molecules. See DNA replication.

Thymine (T) A nitrogenous base, one member of the base pair AT (adenine and thymine).

Transcription The synthesis of an RNA copy from a sequence of DNA (a gene); the first step in gene expression. Compare translation.

Transfer RNA (tRNA) A class of RNA having structures with triplet nucleotide sequences that are complementary to the triplet nucleotide coding sequences of mRNA. The role of tRNAs in protein synthesis is to bond with amino acids and transfer them to the ribosomes, where proteins are assembled according to the genetic code carried by mRNA.

Transformation A process by which the genetic material carried by an individual cell is altered by incorporation of exogenous DNA into its genome.

Translation The process in which the genetic code carried by mRNA directs the synthesis of proteins from amino acids. Compare transcription.

U

tRNA See transfer RNA.

Uracil A nitrogenous base normally found in RNA but not DNA; uracil is capable of forming a base pair with adenine.

V

Vector See cloning vector.

Virus A noncellular biological entity that can reproduce only within a host cell. Viruses consist of nucleic acid covered by protein; some animal viruses are also surrounded by membrane. Inside the infected cell, the virus uses the synthetic capability of the host to produce progeny viruses.

Y

YAC See yeast artificial chromosome.

Yeast artificial chromosome (YAC) A vector used to clone DNA fragments (up to 400 kb); it is constructed from the telomeric, centromeric, and replication origin sequences needed for replication in yeast cells. Compare cloning vector.

JGI
JOINT GENOME INSTITUTE